Journal of Pure and Applied Algebra 174 (2002) 95–115

JOURNAL OF PURE AND APPLIED ALGEBRA

www.elsevier.com/locate/jpaa

Strong hypergroups of order three

N.J. Wildberger

School of Mathematics, UNSW, 2052 Sydney NSW, Australia

Received 15 January 2001; received in revised form 1 November 2001
Communicated by G.M. Kelly

Abstract

This paper investigates the question of when a finite hypergroup with three elements is *strong*, that is satisfies the condition that its dual signed hypergroup is actually a hypergroup. We classify hermitian hypergroups of order three by weight into two-dimensional families and show that the algebraic conditions arising from duality yield four interesting curves in the plane which bound the character values of strong and non-strong hypergroups. By analysing the relations between these curves we discover that the stratum of strong hypergroups is connected for all weights in the range $[4, \infty)$ except for the subinterval $[5, 5\frac{1}{16}]$ where there are two components. For weight equal to 5 the second component degenerates to a single point, the Golden hypergroup. © 2002 Elsevier Science B.V. All rights reserved.

MSC: 20N20

1. Introduction

The theory of hypergroups was introduced into harmonic analysis in the 70's by the papers of Dunkl [6], Jewett [9], and Spector [14]. Earlier work in the direction of discrete or finite hypergroups is implicit in Frobenius [7] (see also Curtis [5] for a discussion of Frobenius' work in this regard), and developed by Kawada [10], and Bose and Mesner [4]. Hypergroups are closely related to other algebraic objects; for example to association schemes (see Bannai and Ito [1]), to hypercomplex systems (see Berezansky and Kalyushnyi [3]) and to the generalized translation operators of Levitan [11]. Surveys of hypergroup theory can be found in Ross [13], Heyer [8], Litvinov [12], Wildberger [17] and in the book of Bloom and Heyer [2].

E-mail address: n.wildberger@unsw.edu.au (N.J. Wildberger).

PII: S0022-4049(02)00016-6

In this paper we study order three hypergroups, which arise naturally as invariants of strongly regular graphs (see Wildberger [18]). They also appear implicitly in the work of Gauss on the cyclotomic constants associated to quadratic residues as described in Wildberger [15].

We say that a hypergroup $\mathscr{K}$ is *strong* if the dual signed hypergroup $\hat{\mathscr{K}}$ is a hypergroup. Dunkl [6] and Jewett [9] give examples of three element hypergroups some of which are strong and some of which are not strong. We are interested in describing within the family of all hypergroups of order three those which are strong. This question is uninteresting for hypergroups of order two which are always self dual (meaning that the dual $\hat{\mathscr{K}}$ is isomorphic to $\mathscr{K}$) and thus strong.

In Section 1 we review some of the basic facts about harmonic analysis on finite commutative hypergroups. If $\mathscr{K}$ is not hermitian, then the situation is also easy since $\mathscr{K}$ is self dual. This case is treated in Section 2.

If $\mathscr{K}$ is hermitian, as in Sections 3 and 4, the situation is more complicated. Classifying the hypergroups themselves is not hard—they form a three parameter family, which is conveniently split into two-dimensional strata by weight. The weight of an element c is defined to be the inverse of the coefficient of the identity c_0 in the product with c^* and the weight of $\mathscr{K}$ is the sum of the weights of its elements. This is the main structural invariant of a finite hypergroup, and for hermitian hypergroups may take on any value in the range $[4,\infty)$. We show that the hypergroup equations lead to interesting algebraic relations between the character values of the hypergroups and the structure constants, and that these relationships can be easily viewed in a two-dimensional picture.

For a given value W of $\omega(\mathscr{K})$ the exact range of character values which correspond to hypergroups and to strong hypergroups are given by certain rather interesting curves f,g,h,k in a designated region of the plane. Roughly speaking we find that as the weight increases an increasing proportion of the stratum of hypergroups of that weight are strong, at least in the somewhat natural coordinates given by the character values.

A curious fact that emerges is that the set of strong hypergroups of fixed weight W is connected for all values of $4 \leqslant W$ except for $5 \leqslant W < 5\frac{1}{16}$ where there is a small connected component separated from the main body. This piece reduces to just one point when $W=5$ at which point we get the *Golden* hypergroup $\mathscr{K}=\{c_0,c_1,c_2\}$ with relations:

$$c_1^2 = \tfrac{1}{2}c_0 + \tfrac{1}{2}c_2,$$

$$c_2^2 = \tfrac{1}{2}c_0 + \tfrac{1}{2}c_1,$$

$$c_1c_2 = \tfrac{1}{2}c_1 + \tfrac{1}{2}c_2.$$

The techniques involve studying the two-dimensional constraints imposed by the positivity of the hypergroup structure constants on the character values, identifying the equations of the boundary curves which result and employing calculus to study their relationships. The convexity of these boundary curves plays an interesting role.

We have included a summary and several figures which illustrate the various geometric relationships involved and show the dependence of the bounding curves on the weight W. To extend this work to an understanding of higher order hypergroups seems a significant challenge.

2. Notation and basic facts

A finite hypergroup is a set $\mathscr{K}=\{c_0,c_1,\ldots,c_n\}$ together with a $*$-algebra structure on the complex vector space $\mathbb{C}\mathscr{K}$ spanned by $\mathscr{K}$ which satisfies some additional axioms. Let the structure equations be

$$c_i c_j = \sum_k n_{ij}^k c_k,$$

where we adopt the convention that summations always range over $\{0,1,\ldots,n\}$. Then we assume the following:

(1) $n_{ij}^k \in \mathbb{R}$ and $n_{ij}^k \geqslant 0\ \forall i,j,k$,
(2) $\sum_k n_{ij}^k = 1\ \forall i,j$,
(3) c_0 is the identity,
(4) $\mathscr{K}^* = \mathscr{K}$ and n_{ij}^0 is non-zero if and only if $c_i^* = c_j$.

If $c_i^* = c_i\ \forall i$, then the hypergroup is called *hermitian*. If $c_i c_j = c_j c_i\ \forall i,j$ then the hypergroup is called *commutative*. Hermitian hypergroups are automatically commutative.

The *order* of $\mathscr{K}=\{c_0,c_1,\ldots,c_n\}$ is $n+1$. Let us write $c_i^* = c_{\sigma(i)}$. The *weight* of c_i is

$$\omega(c_i) = (n_{i\sigma(i)}^0)^{-1} > 0.$$

Note that $\omega(c_0)=1$. Define the *weight* of $\mathscr{K}$ to be

$$\omega(\mathscr{K}) = \sum_k \omega(c_k).$$

Proposition 2.1. *A finite hypergroup $\mathscr{K}$ with three elements is commutative and for any $c_i \in \mathscr{K}$, $\omega(c_i) = \omega(c_i^*)$.*

Proof. This is clear if $\mathscr{K}=\{c_0,c_1,c_2\}$ is hermitian. Otherwise we must have $c_1^* = c_2$ and $c_2^* = c_1$ so we may write $\mathscr{K}=\{c_0,c,c^*\}$. Then

$$cc^* = \alpha c_0 + \beta c + \gamma c^*$$

and

$$cc^* = (cc^*)^* = \alpha c_0 + \beta c^* + \gamma c$$

shows that $\beta = \gamma$. Thus

$$cc^* = \alpha c_0 + \beta c + \beta c^* = \alpha c_0 + \frac{1-\alpha}{2} c + \frac{1-\alpha}{2} c^*$$

as cc^* is a probability measure. The same argument shows that

$$c^*c = \alpha' c_0 + \frac{1-\alpha'}{2} c + \frac{1-\alpha'}{2} c^*.$$

Now we compare

$$(cc^*)c = \alpha c + \frac{1-\alpha}{2} c^2 + \frac{1-\alpha}{2} c^* c$$

and

$$c(c^*c) = \alpha' c + \frac{1-\alpha'}{2} c^2 + \frac{1-\alpha'}{2} cc^*.$$

The coefficient of c_0 in each is, respectively, $((1-\alpha)/2)\alpha'$ and $((1-\alpha')/2)\alpha$ and for these to be equal we must have $\alpha = \alpha'$. Thus $\mathscr{K}$ is commutative and $\omega(c) = \omega(c^*)$. $\square$

Now we recount some basic facts about harmonic analysis on a finite commutative hypergroup $\mathscr{K} = \{c_0, c_1, \ldots, c_n\}$, see [16] for details. A *character* of $\mathscr{K}$ is a mapping $\chi : \mathscr{K} \to \mathbb{C}$ such that

$$\chi(c_i)\chi(c_j) = \sum_k n_{ij}^k \chi(c_k) \quad \forall i, j.$$

We let $\hat{\mathscr{K}}$ denote the set of characters of $\mathscr{K}$. A crucial fact is that the algebra $\mathbb{C}\mathscr{K}$ is semisimple. This implies that there is a basis $\{e_0, e_1, \ldots, e_n\}$ of $\mathbb{C}\mathscr{K}$ for which the operators of multiplication by c_i are simultaneously diagonal, that is

$$c_i e_j = \chi_j(c_i) e_j \tag{2.1}$$

for some functions $\chi_j : \mathscr{K} \to \mathbb{C}$. Furthermore we may normalize the e_i's to be orthogonal idempotents, that is $e_i e_j = \delta_{ij} e_i$. The χ_j's are then exactly the characters of $\mathscr{K}$, and have the additional property that

$$\chi_j(c_i^*) = \overline{\chi_j(c_i)} \quad \forall i, j.$$

If we set $\chi_0 = 1$, then the corresponding idempotent

$$e_0 = \frac{1}{\omega(\mathscr{K})} \sum_k \omega(c_k) c_k$$

is called the *Haar measure* of $\mathscr{K}$; it satisfies

$$c_i e_0 = e_0 \quad \forall i.$$

Let $\mathscr{F}(\mathscr{K})$ be the space of complex valued functions on $\mathscr{K}$ with the inner product

$$\langle f, g \rangle = \omega(\mathscr{K})^{-1} \sum_i \omega(c_i) f(c_i) \overline{g(c_i)}.$$

Then $\hat{\mathscr{K}} = \{\chi_0, \chi_1, \ldots, \chi_n\}$ is an orthogonal basis of $F(\mathscr{K})$. We may consider the pointwise product of characters and thus write

$$\chi_i \chi_j = \sum_k m_{ij}^k \chi_k \quad \text{and} \quad \chi_i^* = \bar{\chi}_i.$$

The dual object $\hat{\mathscr{K}}$ is not necessarily a hypergroup under this multiplication but the only axiom that can fail is A1. There is a theory of duality which takes this fact into account and involves the more general notion of signed hypergroup—see Wildberger [16]. If $\hat{\mathscr{K}}$ is a hypergroup, then we say $\mathscr{K}$ is *strong*.

3. The non-hermitian case

Suppose $\mathscr{K}$ is not hermitian so that we may write $\mathscr{K} = \{c_0, c, c^*\}$. Then the argument of the proof of Proposition 2.1 shows that we may write

$$cc^* = \alpha c_0 + \frac{1-\alpha}{2} c + \frac{1-\alpha}{2} c^*,$$

$$cc = \gamma c + (1-\gamma) c^*$$

and therefore

$$c(cc^*) = \alpha c + \frac{1-\alpha}{2} c^2 + \frac{1-\alpha}{2} cc^*$$

$$(cc)c^* = \gamma cc^* + (1-\gamma) c^* c^*.$$

Considering the coefficient of c_0 in both of these last two equations, we see that $\alpha(1-\alpha)/2 = \gamma\alpha$ which gives $\gamma = (1-\alpha)/2$. Thus

$$cc = \frac{1-\alpha}{2} c + \frac{1+\alpha}{2} c^*$$

for some $0 < \alpha \leqslant 1$. Note that if $\alpha = 1$, we get the group $\mathbb{Z}_3$.

The character table of $\mathscr{K}$ has the form

	c_0	c	c^*
χ_0	1	1	1
χ	1	z	$\bar{z}$
$\bar{\chi}$	1	$\bar{z}$	z

Then

$$z^2 = \chi(c)^2 = \frac{1-\alpha}{2} \chi(c) + \frac{1+\alpha}{2} \overline{\chi(c)} = \frac{1-\alpha}{2} z + \frac{1+\alpha}{2} \bar{z}$$

and writing $z = a + \mathrm{i}b$ and solving for a, b gives

$$z = \frac{-\alpha \pm \mathrm{i}\sqrt{\alpha^2 + 2\alpha}}{2}.$$

Note that the dual $\hat{\mathscr{K}}$ is isomorphic to $\mathscr{K}$. In particular all non-hermitian hypergroups of order three are strong.

4. The hermitian case

Suppose $\mathscr{K}=\{c_0,c_1,c_2\}$ is hermitian and $\omega(c_i)=\omega_i,\ i=1,2$. The structure equations of $\mathscr{K}$ can be written

$$c_1^2=\frac{1}{\omega_1}c_0+\alpha_1c_1+\beta_1c_2, \tag{4.1}$$

$$c_2^2=\frac{1}{\omega_2}c_0+\beta_2c_1+\alpha_2c_2, \tag{4.2}$$

$$c_1c_2=\gamma_1c_1+\gamma_2c_2 \tag{4.3}$$

for some non-negative constants $\alpha_i,\ \beta_i$ and $\gamma_i,\ i=1,2$. There are of course relations among these coefficients; for example by examining in two different ways the coefficient of c_0 in both $c_1^2c_2$ and $c_1c_2^2$ we get

$$\beta_1=\frac{\gamma_1\omega_2}{\omega_1},$$

$$\beta_2=\frac{\gamma_2\omega_1}{\omega_2}.$$

Thus

$$\alpha_1=1-\frac{(1+\gamma_1\omega_2)}{\omega_1},$$

$$\alpha_2=1-\frac{(1+\gamma_2\omega_1)}{\omega_2}.$$

All other relations between the coefficients due to associativity are a consequence of the above. Since $\gamma_1+\gamma_2=1$, it follows that the three quantities $\gamma_1,\ \omega_1$ and ω_2 determine the other coefficients. The conditions $0\leqslant\gamma_2,\ 0\leqslant\alpha_1$ and $0\leqslant\alpha_2$ show that $\gamma_1\leqslant 1$ and that

$$1+\gamma_1\omega_2\leqslant\omega_1, \tag{4.4}$$

$$1+(1-\gamma_1)\omega_1\leqslant\omega_2. \tag{4.5}$$

Proposition 4.1. *Any hermitian hypergroup $\mathscr{K}$ of order three is determined uniquely by the three quantities $0\leqslant\gamma_1\leqslant 1,\ \omega_1\geqslant 1,\ \omega_2\geqslant 1$ subject to the inequalities* (4.4) *and* (4.5). *Furthermore $\omega(\mathscr{K})\geqslant 4$.*

Proof. It suffices to check the final inequality. By adding (4.4) and (4.5) we get

$$\omega(\mathscr{K}) = 1 + \omega_1 + \omega_2 \geqslant 3 + \gamma_1\omega_2 + \gamma_2\omega_1 \geqslant 3 + \gamma_1 + \gamma_2 = 4. \qquad \square$$

Note that the extreme case $\omega(\mathscr{K}) = 4$ can only occur if one of the γ_i is equal to 0 which then forces one of the ω_i to equal 1 and the other to equal 2. Up to isomorphism there is only one such hypergroup given by the equations

$$c_1^2 = c_0,$$

$$c_2^2 = \tfrac{1}{2}c_0 + \tfrac{1}{2}c_1,$$

$$c_1c_2 = c_2.$$

This happens to be the normalized Bose–Mesner algebra of the square viewed as a strongly-regular graph. Now returning to the general hermitian case, to obtain the characters of $\mathscr{K}$ we may treat the c_i as complex numbers, set $c_0 = 1$ and solve for c_i. We obtain

$$c_1^3 - c_1^2(\gamma_2 + \alpha_1) + c_1\left(\gamma_2\alpha_1 - \frac{1}{\omega_1}\right) + \frac{\gamma_2}{\omega_1} - \beta_1 = 0$$

from which we find $c_1 = 1$ or

$$c_1^2 + c_1(\gamma_1 - \alpha_1) - \frac{\gamma_2}{\omega_1} = 0$$

with solutions

$$x = \frac{\alpha_1 - \gamma_1}{2} + \frac{1}{2\omega_1}D \geqslant 0,$$

$$y = \frac{\alpha_1 - \gamma_1}{2} - \frac{1}{2\omega_1}D \leqslant 0,$$

where

$$D = \sqrt{(1 + \gamma_1\omega_2 - \gamma_2\omega_1)^2 + 4\gamma_2\omega_1}.$$

Similarly c_2 satisfies $c_2 = 1$ or

$$c_2^2 + c_2(\gamma_2 - \alpha_2) - \frac{\varepsilon_1}{\omega_2} = 0$$

with solutions

$$z = \frac{\alpha_2 - \gamma_2}{2} - \frac{1}{2\omega_2}D \leqslant 0,$$

$$v = \frac{\alpha_2 - \gamma_2}{2} + \frac{1}{2w_2}D \geqslant 0.$$

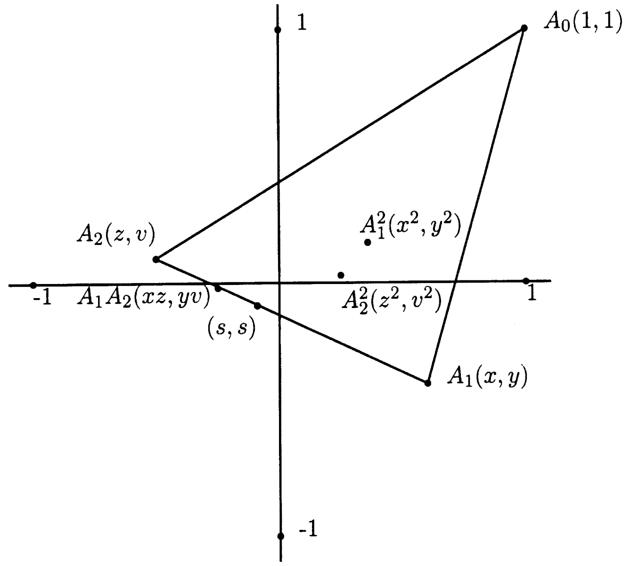

Fig. 1.

Thus the character table of $\mathscr{K}$ has the form

$$\begin{array}{c|ccc} & c_0 & c_1 & c_2 \\ \hline \chi_0 & 1 & 1 & 1 \\ \chi_1 & 1 & x & z \\ \chi_2 & 1 & y & v \end{array}.$$

Orthogonality of its rows yields

$$1+\omega_1 x+\omega_2 z=0,$$

$$1+\omega_1 y+\omega_2 v=0,$$

$$1+\omega_1 xy+\omega_2 zv=0.$$

The character table allows us to concretely identify c_i with the ith column regarded as an element of the algebra $\mathbb{R}^3$. Since any power of c_i is in the convex hull of $\{c_0,c_1,c_2\}$, each of x, y, z and v have absolute value less than or equal to 1. It is useful to consider the triangle in the plane with vertices $A_0(1,1)$, $A_1(x,y)$ and $A_2(z,v)$ as in Fig. 1.

Since $c_1c_2=e_1c_1+e_2c_2$ we see that the point (xz,yv), which we refer to as A_1A_2, lies on the line segment $\overline{A_1A_2}$. Equating slopes on this line gives

$$\frac{v-yv}{z-xz}=\frac{y-yv}{x-xz}=\frac{v-y}{z-x}. \tag{4.6}$$

From this we may infer the relations

$$vx(1-y)(1-z)=yz(1-v)(1-x),$$

$$v(1-y)(z-x)=z(1-x)(v-y),$$

$$y(1-v)(z-x)=x(1-z)(v-y).$$

Now the line segment $\overline{A_1A_2}$ contains a unique point of the form (s,s). Note that $s<0$ since A_1A_2 lies in the third quadrant. Then again consideration of slopes gives

$$\frac{s-v}{s-z}=\frac{s-y}{s-x}$$

from which we find that

$$s=\frac{vx-zy}{v+x-y-z}.$$

Lemma 4.1. $s=-\frac{1}{\omega_1+\omega_2}=-\frac{1}{W-1}$. *In particular* $-\frac{1}{3}\leqslant s$.

Proof. Recall that the element

$$e_0=\frac{1}{W}(c_0+\omega_1c_1+\omega_2c_2)$$

satisfies $c_ie_0=e_0\ \forall c_i\in\mathscr{K}$. It follows that under the identification of the real span of $\mathscr{K}$ with $\mathbb{R}^3$ given by the character table, e_0 must correspond to the vector $(1,0,0)$. Thus $\omega_1c_1+\omega_2c_2$ corresponds to $(\omega_1+\omega_2,-1,-1)$ so that $(\omega_1/(\omega_1+\omega_2))c_1+(\omega_2/(\omega_1+\omega_2))c_2$ corresponds to $(1,-1/(\omega_1+\omega_2),-1/(\omega_1+\omega_2))$. This means that in Fig. 1, $(-1/(\omega_1+\omega_2),-1/(\omega_1+\omega_2))$ is the point of balance on the line segment $\overline{A_1A_2}$ weighted with weights ω_1 and ω_2 on A_1 and A_2, respectively, so that in particular $s=-1/(\omega_1+\omega_2)$. The final inequality follows from the fact that $W\geqslant 4$ as shown in Proposition 4.1. $\square$

Note that

$$D=zy-vx$$

is the area of the triangle ΔOA_1A_2.

If we use (4.6) and solve for z, we get

$$z=\frac{y-1}{yr-1},\tag{4.7}$$

where

$$r=\frac{v+x-1}{vx}.$$

Solving (4.6) for v we get

$$v=\frac{x-1}{xr-1}\tag{4.8}$$

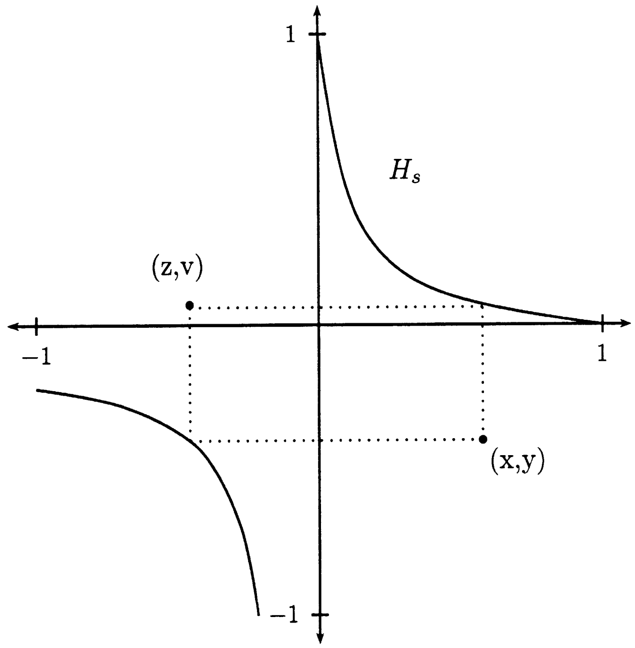

Fig. 2.

and

$$r = \frac{y+z-1}{yz}.$$

Combining the above expressions for r, we see that

$$r = \frac{v+x-y-z}{vx-yz} = \frac{1}{s}.$$

We may thus rewrite (4.7) and (4.8) as

$$z = \frac{s(y-1)}{y-s}, \tag{4.9}$$

$$v = \frac{s(x-1)}{x-s}. \tag{4.10}$$

If we introduce the hyperbola H_s in the (x_1,x_2) plane with equation

$$x_2 = \frac{s(x_1-1)}{x_1-s},$$

then we see that (z,v) is the reflection in H_s of (x,y) as illustrated in Fig. 2.

Let us note some facts about the hyperbola H_s. The top branch goes through $(0,1)$ and $(1,0)$, while the bottom branch goes through $(-1,2s/(1+s))$ and $(2s/(1+s),-1)$. Rewriting the equation of H_s as $(x_1-s)(x_2-s)=s^2-s$ shows that in fact H_s has centre (s,s) and symmetry about the line $x_1=x_2$, which intersects H_s at the points where $x_1=x_2=s\pm\sqrt{s^2-s}$.

5. Character values for hypergroups and strong hypergroups

Our aim is now to find the possible character values $(x, y) \in [0, 1] \times [-1, 0]$ for hermitian hypergroups of a fixed weight W and to discover which of these values correspond to strong hypergroups. Recall by Lemma 4.1 that fixing W is equivalent to fixing s.

It will be convenient to introduce some more notation. Let

$$\alpha = \sqrt{s^2 - s},$$

$$\beta = s^3 - s.$$

Define

$$\tau(t) = \frac{s(t-1)}{t-s}.$$

Then τ is a bijection on each of the intervals $[0, 1]$ and $[-1, 2s/(1+s)]$, reverses order on each, and has fixed points

$$r_1 = \alpha + s$$

and

$$r_2 = \alpha - s.$$

The hypergroup condition tells us that the points $A_1^2(x^2, y^2)$ and $A_2^2(z^2, v^2)$ both should lie inside or on the triangle $\Delta A_0A_1A_2$. The point $A_1^2(x^2, y^2)$ lies inside the triangle $\Delta A_0A_1A_2$ if and only if

$$\frac{1-y^2}{1-x^2} \geqslant \frac{1-v}{1-z}.$$

Using (4.9) and (4.10), this becomes the condition that (x, y) lies above the curve

$$xy(x+y) - s(x^2+xy+y^2) + s = 0. \tag{5.1}$$

The point $A_2^2(z^2, v^2)$ lies in the triangle $\Delta A_0A_1A_2$ if and only if

$$\frac{1-z^2}{1-v^2} \leqslant \frac{1-y}{1-x}.$$

In terms of x and y, a straightforward calculation show that this is the condition that (x, y) lies below the curve

$$y^2x^2 + s(x^2y^2 - 2x^2y - 2y^2x) + s^2(x^2 - 2yx^2 - x + 7yx - 2xy^2 + y^2 - y) \\ + s^3(x^2 - 3x + xy + 2 - 3y + y^2) = 0. \tag{5.2}$$

Theorem 5.1. *The curve* (5.2) *in the region* $[0, 1] \times [-1, 0]$ *is given by the function*

$$f(x) = -\frac{x}{2} - \frac{1}{2}\frac{(x^3 + 3sx^2 - 4s)^{1/2}}{(x-s)^{1/2}}.$$

This function on the interval $[0,1]$ *satisfies the following*:

(1) $f(0)=f(1)=-1$,

(2) f *is convex*,

(3) f *takes on its maximum at* $x=x_m$, *the unique root of* $2x^3-3sx^2+s$ *in* $[0,1]$, *and* $f(x_m)=y_m=-2x_m$,

(4) $x_m > s+\alpha$,

(5) $y_m < s$.

Proof. Since (5.2) is quadratic in y, we may solve it to obtain

$$y=-\frac{x}{2}\pm\frac{1}{2}\frac{(x^3+3sx^2-4s)^{1/2}}{(x-s)^{1/2}}.$$

The larger root is easily shown to be non-negative for $0\leqslant x\leqslant 1$, so we obtain $f(x)$ as stated, and (1) is immediate.

(2) Calculation shows that

$$f'(x)=-\frac{1}{2}-\frac{1}{2}\frac{(x^3-3s^2x+2s)}{(x-s)^{3/2}(x^3+3sx^2-4x)^{1/2}},$$

$$f''(x)=\frac{6s(1-s^2)(x^3-s)}{(x-s)^{5/2}(x^3+3sx^2-4s)^{3/2}}.$$

Note that for $-\frac{1}{3}\leqslant s<0$ and $0\leqslant x\leqslant 1$, the terms x^3+3sx^2-4s, $x-s$, $1-s^2$ and x^3-s are all positive. We conclude that f is convex throughout the interval.

(3) Solving $f'(x)=0$ we obtain

$$2x^3-3sx^2+s=0,$$

which has a unique root, say $x=x_m$, in $[0,1]$. On the other hand differentiating (5.2) implicitly and setting $y'=0$ yields the relation $2x+y=0$ so that the unique maximum occurs at the point $(x_m,-2x_m)$.

(4) Since $s<0$, the quantity $2x^3-3sx^2+s=q(x)$ increases as x increases in the interval $[0,1]$. It thus suffices to show that $q(s+\alpha)<0$. Now

$$\begin{aligned} q(s+\alpha) &= (s+\alpha)^2(2\alpha-s)+s \\ &= s(s-1)(2\alpha+2s-1) \end{aligned}$$

is negative if and only if $2\alpha<1-2s$ (square both sides).

(5) From (4) we have $y_m<-2(s+\alpha)$. Thus

$$\begin{aligned} y_m<s &\Leftrightarrow -2\alpha<3s \\ &\Leftrightarrow 4(s^2-s)>9s^2 \\ &\Leftrightarrow -\tfrac{4}{5}<s \end{aligned}$$

which is true by Lemma 4.1. $\square$

We will now identify the function f and its graph in the region $[0,1]\times[-1,0]$. The curve (5.2) can be seen to be obtained from (5.1) by replacing x and y with $\tau(x)$ and $\tau(y)$, respectively.

Thus we define

$$g = \{(\tau(x), \tau(y)) \mid (x,y) \in f\}.$$

Let us also define the curves

$$h = \{(\tau(x), y) \mid (x,y) \in f\},$$

$$k = \{(x, \tau(y)) \mid (x,y) \in f\}.$$

The four curves f, g, h, k can be identified with the functions f, $\tau \circ f \circ \tau$, $f \circ \tau$, and $\tau \circ f$, respectively. The set of possible character values (x, y) of $\mathscr{K}$ is then given by those points lying above f and below g in the square $[0,1]\times[-1,0]$. By considering the transpose of the character table, we see that $\mathscr{K}$ is strong if in addition (x, y) lies above h and below k. The relative positions of the four curves f, g, h and k (all of which depend on s) will thus be crucial in what follows.

In order to investigate these relations, we introduce new co-ordinates $X_1 = x_1 - s$, $X_2 = x_2 - s$ and define $X = x - s$, $Y = y - s$, $Z = z - s$, and $V = v - s$. The origin is now at the center of the hyperbola H_s which has equation $X_1 X_2 = \alpha^2$. Define

$$\bar{\tau}(X) = \frac{\alpha^2}{X}.$$

Note that $\bar{\tau}$ is a bijection on each of $[-s, 1-s]$ and $[-1-s, (s-s^2)/(1-s)]$, reverses order on each, and has fixed points α and $-\alpha$ on each, respectively.

In the new co-ordinates the curve f becomes the curve F with equation

$$XY(X + Y + 3s) - \beta = 0,$$

the curve g becomes the curve G with equation

$$\alpha^4(X + Y) + 3s\alpha^2 XY - (s+1)X^2Y^2 = 0,$$

the curve h becomes the curve H with equation

$$\alpha^2 Y + XY^2 + 3XYs - (s+1)X^2 = 0$$

and the curve k becomes the curve K with equation

$$\alpha^2 X + YX^2 + 3XYs - (s+1)Y^2 = 0.$$

These are all quadratic in Y so we may find explicit forms for these curves in the regions we need:

$$F(X) = -\frac{1}{2}\frac{X^2 + 3sX + \sqrt{X^4 + 6sX^3 + 9s^2X^2 + 4s^3X - 4sX}}{X},$$

$$G(X) = \frac{s}{2}\{s^4 - 3s^3 + 3s^3X + 3s^2 - 6s^2X - s + 3sX$$

$$+(s-1)^2\sqrt{s^4+6s^3X-2s^3+9s^2X^2-6s^2X+s^2+4sX^3+4X^3}\}/(s^2-1)X^2,$$

$$H(X) = -\frac{1}{2}\frac{s^2+3sX-s+\sqrt{s^4+6s^3X-2s^3+9s^2X^2-6s^2X+s^2+4sX^3+4X^3}}{X},$$

$$K(X) = \frac{1}{2}\frac{sX^2+3s^2X-X^2-3sX+(s-1)^2\sqrt{X(4s^3+9s^2X+6sX^2-4s+X^3)}}{(s^2-1)}.$$

From Theorem 5.1(3) we find that the maximum of F occurs at the point (X_m, Y_m) satisfying $2X_m + Y_m + 3s = 0$ and $-X_m^2 Y_m = \beta$, and furthermore by (4) we see that $X_m > \alpha$.

Lemma 5.1. *F is below the line $Y = -\alpha \Leftrightarrow -\frac{16}{65} < s$.*

Proof. Since $X_m = \frac{1}{2}(-Y_m - 3s)$, $Y_m < -a \Leftrightarrow X_m > \frac{1}{2}(\alpha - 3s)$. As Y_m increases, $-Y_m X_m^2$ decreases so that

$$\begin{aligned} Y_m < -\alpha &\Leftrightarrow -Y_m X_m^2 > \alpha\left(\frac{\alpha - 3s}{2}\right)^2 \\ &\Leftrightarrow 4(s^3 - s) > \alpha(\alpha^2 - 6s\alpha + 9s^2) \\ &\Leftrightarrow 10s^3 - 6s^2 - 4s > \alpha(10s^2 - s) \\ &\Leftrightarrow 2(5s^2 - 3s - 2) < \alpha(10s - 1). \end{aligned}$$

Because $-\frac{1}{3} \leqslant s$, both sides of the latter expression are negative. When we square both sides and simplify, we obtain

$$Y_m < -\alpha \quad \Leftrightarrow \quad 65s^2 - 49s - 16 < 0.$$

This last quadratic has roots 1 and $-\frac{16}{65}$ from which the claim follows. □

Lemma 5.2. *For any s the curves F and G intersect in at most two places.*

Proof. From the equation for F, we may write

$$X + Y = \frac{\beta}{XY} - 3s. \tag{5.3}$$

Substituting this into the equation for G, we obtain

$$\alpha^4\beta - 3s\alpha^4(XY) + 3s\alpha^2(XY)^2 - (s+1)(XY)^3 = 0.$$

This cubic in XY has at most three roots, but in fact since the product of the roots is $\alpha^4\beta/(s+1) > 0$, there are at most two negative roots. Now from (5.3), XY determines $X + Y$ and so X and Y up to permutation, but since $X > 0$ and $Y < 0$, we see there are at most two points of intersection. □

Clearly if we knew that the curve G is concave in the interval $[-s, 1-s]$, then we could deduce Lemma 5.2 immediately. It is somewhat curious that while G often appears concave upon casual inspection, it is generally not so.

Lemma 5.3. *If F, G intersect in points $(X_1, Y_1), (X_2, Y_2)$ then $\bar{\tau}(X_1) = X_2$, $\bar{\tau}(Y_1) = Y_2$.*

Proof. This follows from the fact that $(X, Y) \in F \Leftrightarrow (\bar{\tau}(X), \bar{\tau}(Y)) \in G$. $\square$

Lemma 5.4. (i) *The curves F and G intersect at two points $\Leftrightarrow$ the point $(\alpha, -\alpha)$ lies under the curve F.*

(ii) *The curves F and G are tangential $\Leftrightarrow$ their common intersection is the point $(\alpha, -\alpha)$.*

Proof. (i) If $(\alpha, -\alpha)$ lies under the curve F, then F has on it a point (α, α') with $\alpha' > -\alpha$ and so G contains the point $(\bar{\tau}(\alpha), \bar{\tau}(\alpha')) = (\alpha, \bar{\tau}(\alpha'))$ where $\bar{\tau}(\alpha') < -\alpha$. Thus, some point of G is under some point of F, and by continuity and the values of F and G at the end points, F and G must intersect in at least two points, and so by Lemma 5.2 in exactly two points.

Conversely suppose F and G intersect exactly at (X_1, Y_1) and (X_2, Y_2) with $X_1 < X_2$. By Lemma 5.3, we must have $\bar{\tau}(X_1) = X_2$ so $X_1 < \alpha < X_2$. Then G must lie under F throughout this interval and in particular if $(\alpha, Y_F) \in F$ and $(\alpha, Y_G) \in G$, then $Y_G < Y_F$ and also $Y_G = \bar{\tau}(Y_F)$. But this means that $Y_G < -\alpha < Y_F$, that is $(\alpha, -\alpha)$ lies under F.

(ii) Obvious. $\square$

Lemma 5.5. *The point $(\alpha, -\alpha)$ lies under the curve $F \Leftrightarrow s < -\frac{1}{4}$.*

Proof. From the defining equation of F, $(\alpha, -\alpha)$ lies under the curve of F

$$\Leftrightarrow \alpha(-\alpha)(\alpha + (-\alpha) + 3s) - \beta > 0$$

$$\Leftrightarrow 4s^3 - 3s^2 - s < 0$$

$$\Leftrightarrow 4s^2 - 3s - 1 > 0.$$

This quadratic has roots 1 and $-\frac{1}{4}$, so the result follows. $\square$

Tangency of F and G happens when $s = -\frac{1}{4}$ and occurs at the point $(\alpha, -\alpha) = (\sqrt{5}/4, -\sqrt{5}/4)$.

The curves H and K given previously may also be described by

$$H = \{(\bar{\tau}(X), Y) \mid (X, Y) \in F\},$$

$$K = \{(X, \bar{\tau}(Y)) \mid (X, Y) \in F\}.$$

Lemma 5.6. *The curves F and H intersect at $X = \alpha$ and $X = 1 - s$. For $\alpha < X < 1 - s$, the curve H lies under the curve F.*

Proof. From the definition of H, the curves intersect at $X = \alpha$ since α is a fixed point of $\bar{\tau}$. They also intersect at $X = 1 - s$ since $\bar{\tau}(1 - s) = -s$ and $F(-s) = F(1 - s) = -1$ from Theorem 5.1. To show that in the given range H lies under F we need show that for any $Y \in [-1 - s, -s]$ for which there exist X_L and X_R with $F(X_L) = F(X_R) = Y$

and $X_L \leqslant X_R$, then $\bar{\tau}(X_L) < X_2$. Since $0 < -s \leqslant X_L$, this inequality is just $\alpha^2 < X_L X_R$. Now recalling the equation of F

$$X^2 Y + X(Y^2 + 3sY) - \beta = 0,$$

we see that for fixed Y the product of the roots $X_L X_R$ is $-\frac{\beta}{Y}$. But by Theorem 5.1(5), we have that $Y < 0$ so that

$$\begin{aligned} \alpha^2 < -\frac{\beta}{Y} &\Leftrightarrow -\frac{\beta}{\alpha^2} < Y \\ &\Leftrightarrow -1 - s < Y \end{aligned}$$

and this is true in the range considered. □

Proposition 5.1. *If $-\frac{16}{65} < s$ then F and K do not intersect and G and H do not intersect. If F and K intersect, they do so on the line $Y = -\alpha$ in exactly 2 places unless $s = -\frac{16}{65}$ in which case they are tangent at the point $(\alpha, -\alpha) = (\frac{36}{\sqrt{65}}, -\frac{36}{\sqrt{65}})$.*

We are now in a position to describe all possible character values for order three hypergroups and for strong order three hypergroups. We restate the results established so far in terms of the original variables x, y, etc. and the original functions f, g, h and k.

First note that if (x, y) is a character value, so is $(\tau(x), \tau(y))$, obtained simply by relabelling c_1 and c_2. We may thus without loss of generality insist that $x \geqslant \alpha + s$. This means the only remaining identifications are along the boundary line $x = \alpha + s$; that is the points $(\alpha + s, y), (\alpha + s, \tau(y))$ correspond to isomorphic hypergroups if they occur.

Let R_s be the set of all points (x, y) such that $x \geqslant \alpha + s$ and such that

	c_0	c_1	c_2
χ_0	1	1	1
χ_1	1	x	$\tau(y)$
χ_2	1	y	$\tau(x)$

is the character table of a hypergroup. Let R_s^+ be the subset of R_s corresponding to the set of strong hypergroups.

Theorem 5.2. *R_s consists of all points (x, y) which are on or above the curve f, on or below the curve g, and to the right of or on the line $x = \alpha + s$. The subset R_s^+ consists of those $(x, y) \in R_s$ which are in addition on or below the curve k.*

Proof. The first statement is a reformulation of the conditions (5.1) and (5.2) while R_s^+ consists of those $(x, y) \in R_s$ such that $(x, \tau(y)) \in R_s$. This amounts to two more equations determined by k and h. However by Lemma 5.6, the condition that (x, y) is above the curve h in the range $x \geqslant \alpha + s$ is a consequence of the fact that it is above the curve f. Thus, the only curves we need to be concerned with are f, g, and k. □

The set R_s as defined above is connected for all possible values of s; i.e. $-\frac{1}{3} < s < 0$. However the set R_s^+ exhibits a somewhat interesting behaviour. As s decreases from 0, R_s^+ is connected until $s = -16/65$, at which point a small triangle is 'pinched off' from the main body. This small triangle decreases in size as s decreases to $s = -1/4$ at which point it is a single point, and vanishes entirely for $-1/3 < s < -1/4$. The limiting point when $s = -1/4$ gives us the *Golden* hypergroup:

$$c_1^2 = \tfrac{1}{2}c_0 + \tfrac{1}{2}c_2,$$

$$c_2^2 = \tfrac{1}{2}c_0 + \tfrac{1}{2}c_1,$$

$$c_1c_2 = \tfrac{1}{2}c_1 + \tfrac{1}{2}c_2.$$

It is not possible to deform the Golden hypergroup within the family of all strong hypergroups to obtain a hypergroup with strictly smaller weight.

6. Summary

The most important invariant of a finite commutative hypergroup $\mathscr{K}$ is its weight $\omega(\mathscr{K}) = W$. Non-hermitian hypergroups $\mathscr{K}$ of order three are determined by a single parameter $0 < \alpha \leqslant 1$ and satisfy $\hat{\mathscr{K}} \simeq \mathscr{K}$. The range of possible weights W is $[3, \infty)$.

The family of hermitian order three hypergroups has three parameters. The weight W of any such hypergroup is constrained to be in the range $[4, \infty)$. The columns of the character table determine three points in the plane; $A_0(1,1)$, $A_1(x, y)$ and $A_2(z, v)$ which we may take to be in first, fourth and second quadrants, respectively. The line segment $\overline{A_1A_2}$ intersects the diagonal in the point (s, s) where $s = -1/(W-1)$. If we know W, and therefore s, we can say that A_2 is determined by A_1 by a reflection in a certain hyperbola H_s. For fixed W, it thus suffices to describe the possible positions for the point A_1 in the square $[0, 1] \times [-1, 0]$.

The condition that $\mathscr{K}$ is a hypergroup requires that both A_1^2 and A_2^2 lie in the triangle $\Delta A_0A_1A_2$. This gives us two bounding constraints on the position of A_1 which determine two curves f and g (both depending on s) in the square $[0, 1] \times [-1, 0]$—the possible positions of A_1 is the set between these curves. If we take into consideration that the choice of the order of columns of the character table was arbitrary, we find that we may restrict the character values to lie on or to the right of a natural line of symmetry. This then gives us an explicit description of all the possible character values of a hypergroup of a given weight.

When we inquire into which of these hypergroups are in addition strong (that is their duals are also hypergroups), we get two additional curves h and k determined by the transpose of the character table. The resulting set $R_s^+ \subset [0, 1] \times [-1, 0]$ which describes all possible positions of A_1 for strong hypergroups of a given weight, displays

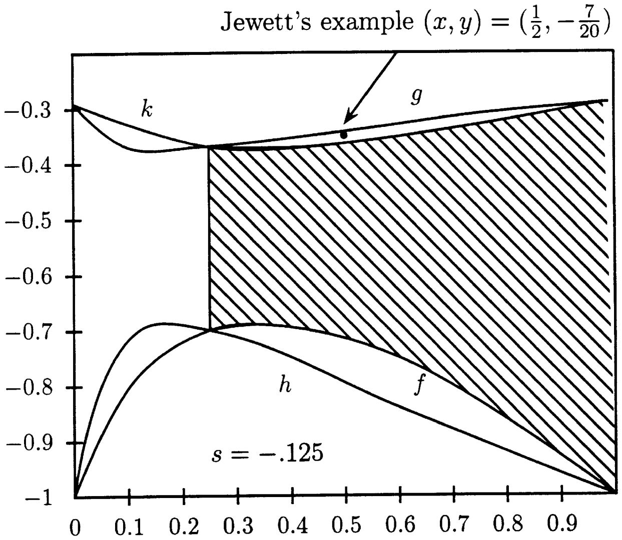

Fig. 3.

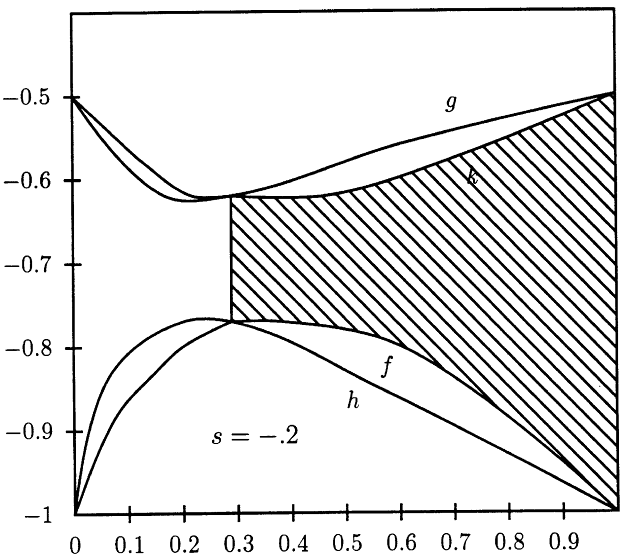

Fig. 4.

an interesting phenomenon—for the values $5\frac{1}{16} \leqslant W$ it is connected, for the values $5 \leqslant W < 5\frac{1}{16}$ however there is an additional component, which decreases in size as W decreases until for $W = 5$ it is a single point and disappears for $4 \leqslant W < 5$.

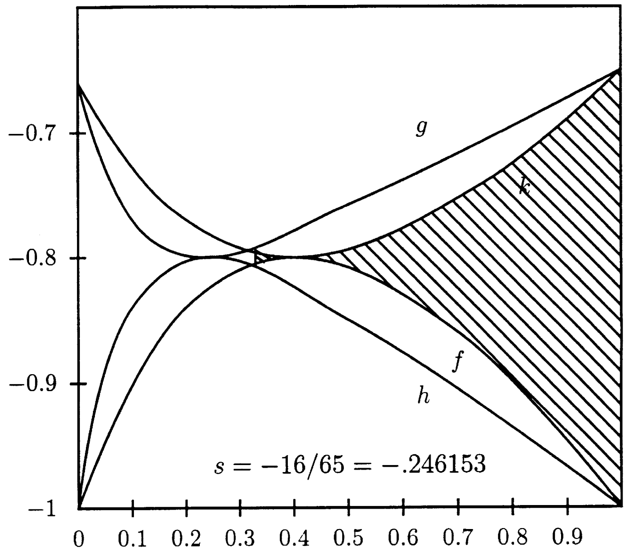

Fig. 5.

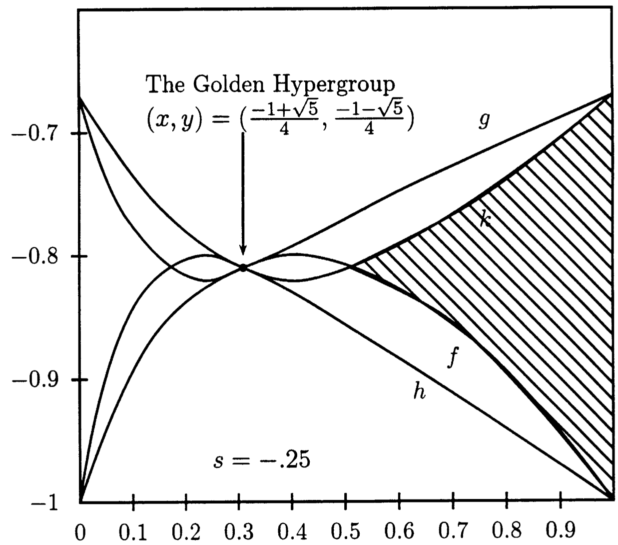

Fig. 6.

The various curves and regions are illustrated in Figs. 3–7 for different values of s (and thus W). Regions shaded signify positions of A_1 which correspond to strong hypergroups. The region between the curves g and k correspond to non-strong hyper-

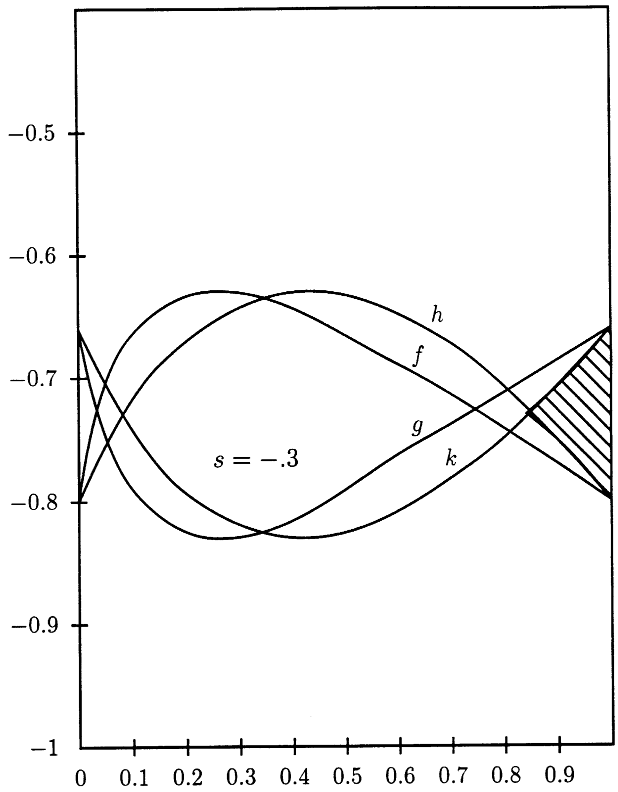

Fig. 7.

groups. We draw attention to Fig. 3 in which we locate the character values of the example of Jewett [9] of a non-strong hypergroup.

Acknowledgements

The author would like to thank the referee for many helpful suggestions and critical comments.

References

[1] E. Bannai, T. Ito, Algebraic Combinatorics I—Association Schemes, Benjamin and Cummings, Menlo Park, 1984.

[2] Y.M. Berezansky, A.A. Kalyushnyi, Hypercomplex systems with locally compact bases, Sel. Math. Sov. 4 (1985) 151–200.

[3] W.R. Bloom, H. Heyer, Harmonic analysis of probability measures on hypergroups, de Gruyter Studies in Mathematics, Vol. 20, Walter de Gruyter, Berlin, 1995.

[4] R.C. Bose, D.M. Mesner, On linear associative algebras corresponding to association schemes of partially balanced designs, Ann. Math. Statist. 30 (1959) 21–38.
[5] C.W. Curtis, Representation theory of finite groups: from Frobenius to Brauer, Math. Intell. 14 (1992) 48–57.
[6] C.F. Dunkl, The measure algebra of a locally compact hypergroup, Trans. Amer. Math. Soc. 179 (1973) 331–348.
[7] G. Frobenius, Über Gruppencharaktere, Gesammelte Abhandlungen, Vol. III, Springer, Berlin, pp. 1–37.
[8] H. Heyer, Probability theory on hypergroups: a survey in: H. Heyer (Ed.), Probability Measures on Groups VII, Lecture Notes in Mathematics, Vol. 1064, Springer, Berlin, 1984, pp. 481–550.
[9] R.I. Jewett, Spaces with an abstract convolution of measures, Adv. Math. 18 (1975) 1–101.
[10] Y. Kawada, Über den dualitätssatz der Charaktere nichtcommutative Gruppen, Proc. Phys. Math. Soc. Japan 24 (3) (1942) 97–109.
[11] B.M. Levitan, Generalized translation operators and some of their applications, Israel Program for Scientific Translations, Jerusalem, 1962.
[12] G.L. Litvinov, Hypergroups and hypergroup algebras, J. Soviet Math. 38 (1987) 1734–1761.
[13] K. Ross, Hypergroups and centers of measure algebras, Symposia Math. 22 (1977) 119–172.
[14] R. Spector, Mesures invariants sur les hypergroups, Trans Amer. Math. Soc. 239 (1978) 147–165.
[15] N.J. Wildberger, Finite commutative hypergroups and applications from group theory to conformal field theory, in: W. Connett, O. Gebuhrer, A. Schwartz (Eds.), Proceedings of Applications of Hypergroups and Related Measure Algebras, Seattle 1993, Contemporary Mathematics, Vol. 183, American Mathematical Society, Providence, RI, 1995, pp. 413–434.
[16] N.J. Wildberger, On the algebraic structure of Gaussian periods, J. Number Theory 57 (2) (1996) 278–291.
[17] N.J. Wildberger, Lagrange's theorem and integrality for finite commutative hypergroups with applications to strongly-regular graphs, J. Algebra. 182 (1996) 1–37.
[18] N.J. Wildberger, Duality and entropy for finite commutative hypergroups and fusion rule algebras, J. London Math. Soc. 56 (2) (1997) 275–291.